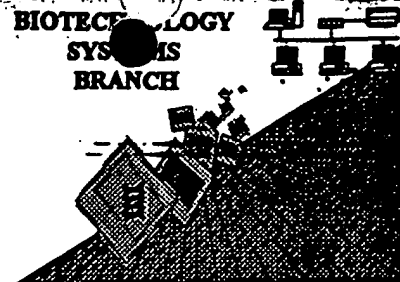


## **RAW SEQUENCE LISTING** **ERROR REPORT**



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/643,755

Source: OIPE

Date Processed by STIC: 8-30-2000

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR FURTHER INFORMATION, PLEASE TELEPHONE MARK SPENCER, 703-308-4212.

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

### **Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

**Checker Version 3.0 can be down loaded from the USPTO website at the following address:**

**<http://www.uspto.gov/web/offices/pac/checker>**

# Sequence Listing Error Summary

## ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/643,755

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 ☐ Wrapped Nucleics      The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2 ☐ Wrapped Aminos      The amino acid number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3 ☐ Incorrect Line Length      The rules require that a line not exceed 72 characters in length. This includes spaces.
- 4 ☐ Misaligned Amino Acid Numbering      The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
- 5 ☐ Non-ASCII      This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6 ☐ Variable Length      Sequence(s) \_\_\_\_\_ contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and indicate in the (b) feature section that some may be missing.
- 7 ☐ PatentIn ver. 2.0 "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s) \_\_\_\_\_. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 8 ☐ Skipped Sequences (OLD RULES)      Sequence(s) \_\_\_\_\_ missing. If intentional, please use the following format for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X:  
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
(x) SEQUENCE DESCRIPTION:SEQ ID NO:X:  
This sequence is intentionally skipped  
  
Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9 ☐ Skipped Sequences (NEW RULES)      Sequence(s) \_\_\_\_\_ missing. If intentional, please use the following format for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 10 ☐ Use of n's or Xaa's (NEW RULES)      Use of n's and/or Xaa's have been detected in the Sequence-Listing.  
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 11 ☐ Use of <213>Organism (NEW RULES)      Sequence(s) \_\_\_\_\_ are missing this mandatory field or its response.
- 12 ☒ Use of <220>Feature (NEW RULES)      Sequence(s) <sup>#4</sup> are missing the <220>Feature and associated headings.  
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"  
Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)
- 13 ☐ PatentIn ver. 2.0 "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other means to copy file to floppy disk.

OIPE

RAW SEQUENCE LISTING                      DATE: 08/30/2000  
 PATENT APPLICATION: US/09/643,755              TIME: 14:32:19

Input Set : A:\Sequence  
 Output Set: N:\CRF3\08302000\I643755.raw

Does Not Comply  
 Corrected Diskette Needed

4 <110> APPLICANT: van Rooijen, Gijs  
 5 Keon, Richard Glenn  
 6 Boothe, Joseph  
 7 Shen, Yin  
 10 <120> TITLE OF INVENTION: Commercial Production of Chymosin in Plants  
 12 <130> FILE REFERENCE: 9369-153  
 C--> 14 <140> CURRENT APPLICATION NUMBER: US/09/643,755  
 C--> 15 <141> CURRENT FILING DATE: 2000-08-23  
 17 <160> NUMBER OF SEQ ID NOS: 4  
 19 <170> SOFTWARE: PatentIn Ver. 2.0  
 21 <210> SEQ ID NO: 1  
 22 <211> LENGTH: 1173  
 23 <212> TYPE: DNA  
 24 <213> ORGANISM: Bovine  
 26 <220> FEATURE:  
 27 <221> NAME/KEY: CDS  
 28 <222> LOCATION: (1)..(1173)  
 30 <400> SEQUENCE: 1  
 31 atg aac ttc ctt aag tct ttc cct ttc tac gct ttc ctt tgt ttc ggt 48  
 32 Met Asn Phe Leu Lys Ser Phe Pro Phe Tyr Ala Phe Leu Cys Phe Gly  
 33 1 5 10 15  
 35 caa tac ttc gtt gct gtt act cac gct gct gag atc acc cgc att cct 96  
 36 Gln Tyr Phe Val Ala Val Thr His Ala Ala Glu Ile Thr Arg Ile Pro  
 37 20 25 30  
 39 ctc tac aaa ggt aag tct ctc cgt aag gcg ctg aag gaa cat gga ctt 144  
 40 Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly Leu  
 41 35 40 45  
 43 cta gaa gac ttc ttg cag aaa caa cag tat ggc atc agc agc aag tac 192  
 44 Leu Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys Tyr  
 45 50 55 60  
 47 tcc ggc ttc ggt gaa gtt gct agc gtg cca ctt acc aac tac ctt gat 240  
 48 Ser Gly Phe Gly Glu Val Ala Ser Val Pro Leu Thr Asn Tyr Leu Asp  
 49 65 70 75 80  
 51 agt caa tac ttt ggg aag atc tac ctc gga acc ccg cct caa gag ttc 288  
 52 Ser Gln Tyr Phe Gly Lys Ile Tyr Leu Gly Thr Pro Pro Gln Glu Phe  
 53 85 90 95  
 55 acc gtt ctc ttt gat act ggt tcc tct gac ttc tgg gtt ccc tct atc 336  
 56 Thr Val Leu Phe Asp Thr Gly Ser Ser Asp Phe Trp Val Pro Ser Ile  
 57 100 105 110  
 59 tac tgc aag agc aat gcc tgc aag aac cac caa aga ttc gat ccg aga 384  
 60 Tyr Cys Lys Ser Asn Ala Cys Lys Asn His Gln Arg Phe Asp Pro Arg  
 61 115 120 125  
 63 aag tcg tcc acc ttc cag aac tta ggc aaa ccc ttg tct ata cac tac 432  
 64 Lys Ser Ser Thr Phe Gln Asn Leu Gly Lys Pro Leu Ser Ile His Tyr  
 65 130 135 140  
 67 ggt aca ggt agc atg caa gga atc tta ggc tat gat acc gtc act gtc 480  
 68 Gly Thr Gly Ser Met Gln Gly Ile Leu Gly Tyr Asp Thr Val Thr Val

see p. 4, 6

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/643,755

DATE: 08/30/2000  
 TIME: 14:32:19

Input Set : A:\Sequence  
 Output Set: N:\CRF3\08302000\I643755.raw

```

69 145          150          155          160
71 tcc aac att gtg gac att caa cag aca gta gga ctt agc acc caa gaa 528
72 Ser Asn Ile Val Asp Ile Gln Gln Thr Val Gly Leu Ser Thr Gln Glu
73          165          170          175
75 cca ggt gat gtc ttc acc tat gca gaa ttc gat ggc atc ctt ggt atg 576
76 Pro Gly Asp Val Phe Thr Tyr Ala Glu Phe Asp Gly Ile Leu Gly Met
77          180          185          190
79 gca tac cca tcg ctc gcg tca gag tac tcg ata cct gtg ttt gac aac 624
80 Ala Tyr Pro Ser Leu Ala Ser Glu Tyr Ser Ile Pro Val Phe Asp Asn
81          195          200          205
83 atg atg aac cga cac cta gta gct caa gac ttg ttc tcg gtt tac atg 672
84 Met Met Asn Arg His Leu Val Ala Gln Asp Leu Phe Ser Val Tyr Met
85          210          215          220
87 gac agg aat ggc cag gag agc atg ctc acg ctt gga gct att gat cca 720
88 Asp Arg Asn Gly Gln Ser Met Leu Thr Leu Gly Ala Ile Asp Pro
89 225          230          235          240
91 tcc tac tac aca gga tct ctt cac tgg gtt cca gtc act gtg cag cag 768
92 Ser Tyr Tyr Thr Gly Ser Leu His Trp Val Pro Val Thr Val Gln Gln
93          245          250          255
95 tac tgg caa ttc act gtg gac agt gtc acc atc agc ggt gtg gtt gtt 816
96 Tyr Trp Gln Phe Thr Val Asp Ser Val Thr Ile Ser Gly Val Val Val
97          260          265          270
99 gca tgt gaa ggt gga tgt caa gct atc ttg gat acc ggt acg tcc aag 864
100 Ala Cys Glu Gly Gly Cys Gln Ala Ile Leu Asp Thr Gly Thr Ser Lys
101          275          280          285
103 ctg gtc gga cct agc agc gac att ctc aac att cag caa gct att gga 912
104 Leu Val Gly Pro Ser Ser Asp Ile Leu Asn Ile Gln Gln Ala Ile Gly
105          290          295          300
107 gcc aca cag aac cag tac ggt gag ttt gac ata gat tgc gac aac ctt 960
108 Ala Thr Gln Asn Gln Tyr Gly Glu Phe Asp Ile Asp Cys Asp Asn Leu
109 305          310          315          320
111 agc tac atg cct aca gtt gtc ttt gag atc aac ggc aag atg tac cca 1008
112 Ser Tyr Met Pro Thr Val Val Phe Glu Ile Asn Gly Lys Met Tyr Pro
113          325          330          335
115 ctg acc ccc tcc gcc tat acc agc cag gat caa ggg ttc tgc acc agt 1056
116 Leu Thr Pro Ser Ala Tyr Thr Ser Gln Asp Gln Gly Phe Cys Thr Ser
117          340          345          350
119 gga ttc cag agt gag aac cat tcc cag aaa tgg atc ttg gga gat gtg 1104
120 Gly Phe Gln Ser Glu Asn His Ser Gln Lys Trp Ile Leu Gly Asp Val
121          355          360          365
123 ttc att cgt gag tac tac agc gtc ttt gac agg gcc aac aac ctc gtt 1152
124 Phe Ile Arg Glu Tyr Tyr Ser Val Phe Asp Arg Ala Asn Asn Leu Val
125          370          375          380
127 ggg cta gct aaa gca atc tga 1173
128 Gly Leu Ala Lys Ala Ile
129 385          390
132 <210> SEQ ID NO: 2
133 <211> LENGTH: 390
134 <212> TYPE: PRT

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/643,755

DATE: 08/30/2000

TIME: 14:32:19

Input Set : A:\Sequence

Output Set: N:\CRF3\08302000\I643755.raw

135 &lt;213&gt; ORGANISM: Bovine

137 &lt;400&gt; SEQUENCE: 2

```

138 Met Asn Phe Leu Lys Ser Phe Pro Phe Tyr Ala Phe Leu Cys Phe Gly
139 1 5 10 15
141 Gln Tyr Phe Val Ala Val Thr His Ala Ala Glu Ile Thr Arg Ile Pro
142 20 25 30
144 Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly Leu
145 35 40 45
147 Leu Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys Tyr
148 50 55 60
150 Ser Gly Phe Gly Glu Val Ala Ser Val Pro Leu Thr Asn Tyr Leu Asp
151 65 70 75 80
153 Ser Gln Tyr Phe Gly Lys Ile Tyr Leu Gly Thr Pro Pro Gln Glu Phe
154 85 90 95
156 Thr Val Leu Phe Asp Thr Gly Ser Ser Asp Phe Trp Val Pro Ser Ile
157 100 105 110
159 Tyr Cys Lys Ser Asn Ala Cys Lys Asn His Gln Arg Phe Asp Pro Arg
160 115 120 125
162 Lys Ser Ser Thr Phe Gln Asn Leu Gly Lys Pro Leu Ser Ile His Tyr
163 130 135 140
165 Gly Thr Gly Ser Met Gln Gly Ile Leu Gly Tyr Asp Thr Val Thr Val
166 145 150 155 160
168 Ser Asn Ile Val Asp Ile Gln Gln Thr Val Gly Leu Ser Thr Gln Glu
169 165 170 175
171 Pro Gly Asp Val Phe Thr Tyr Ala Glu Phe Asp Gly Ile Leu Gly Met
172 180 185 190
174 Ala Tyr Pro Ser Leu Ala Ser Glu Tyr Ser Ile Pro Val Phe Asp Asn
175 195 200 205
177 Met Met Asn Arg His Leu Val Ala Gln Asp Leu Phe Ser Val Tyr Met
178 210 215 220
180 Asp Arg Asn Gly Gln Glu Ser Met Leu Thr Leu Gly Ala Ile Asp Pro
181 225 230 235 240
183 Ser Tyr Tyr Thr Gly Ser Leu His Trp Val Pro Val Thr Val Gln Gln
184 245 250 255
186 Tyr Trp Gln Phe Thr Val Asp Ser Val Thr Ile Ser Gly Val Val Val
187 260 265 270
189 Ala Cys Glu Gly Cys Gln Ala Ile Leu Asp Thr Gly Thr Ser Lys
190 275 280 285
192 Leu Val Gly Pro Ser Ser Asp Ile Leu Asn Ile Gln Gln Ala Ile Gly
193 290 295 300
195 Ala Thr Gln Asn Gln Tyr Gly Glu Phe Asp Ile Asp Cys Asp Asn Leu
196 305 310 315 320
198 Ser Tyr Met Pro Thr Val Val Phe Glu Ile Asn Gly Lys Met Tyr Pro
199 325 330 335
201 Leu Thr Pro Ser Ala Tyr Thr Ser Gln Asp Gln Gly Phe Cys Thr Ser
202 340 345 350
204 Gly Phe Gln Ser Glu Asn His Ser Gln Lys Trp Ile Leu Gly Asp Val
205 355 360 365
207 Phe Ile Arg Glu Tyr Tyr Ser Val Phe Asp Arg Ala Asn Asn Leu Val

```

8/30/00

291		35					40				45								
293	gaa	gac	ttc	ttg	cag	aaa	caa	cag	tat	ggc	atc	agc	agc	aag	tac	tcc			1748
294	Glu	Asp	Phe	Leu	Gln	Lys	Gln	Gln	Tyr	Gly	Ile	Ser	Ser	Lys	Tyr	Ser			
295	50					55					60					65			
297	ggc	ttc	ggt	gaa	gtt	gct	agg	gtg	cca	ctt	acc	aac	tac	ctt	gat	agt			1796
298	Gly	Phe	Gly	Glu	Val	Ala	Ser	Val	Pro	Leu	Thr	Asn	Tyr	Leu	Asp	Ser			
299					70					75					80				
301	caa	tac	ttt	ggg	aag	atc	tac	ctc	gga	acc	ccg	cct	caa	gag	ttc	acc			1844
302	Gln	Tyr	Phe	Gly	Lys	Ile	Tyr	Leu	Gly	Thr	Pro	Pro	Gln	Glu	Phe	Thr			
303				85						90				95					
305	gtt	ctc	ttt	gat	act	ggt	tcc	tct	gac	ttc	tgg	gtt	ccc	tct	atc	tac			1892
306	Val	Leu	Phe	Asp	Thr	Gly	Ser	Ser	Asp	Phe	Trp	Val	Pro	Ser	Ile	Tyr			
307			100					105					110						
309	tgc	aag	agc	aat	gcc	tgc	aag	aac	cac	caa	aga	ttc	gat	ccg	aga	aag			1940
310	Cys	Lys	Ser	Asn	Ala	Cys	Lys	Asn	His	Gln	Arg	Phe	Asp	Pro	Arg	Lys			
311		115				120					125								
313	tcg	tcc	acc	ttc	cag	aac	tta	ggc	aaa	ccc	ttg	tct	ata	cac	tac	ggt			1988
314	Ser	Ser	Thr	Phe	Gln	Asn	Leu	Gly	Lys	Pro	Leu	Ser	Ile	His	Tyr	Gly			
315	130				135					140					145				
317	aca	ggt	agc	atg	caa	gga	atc	tta	ggc	tat	gat	acc	gtc	act	gtc	tcc			2036
318	Thr	Gly	Ser	Met	Gln	Gly	Ile	Leu	Gly	Tyr	Asp	Thr	Val	Thr	Val	Ser			
319				150						155				160					
321	aac	att	gtg	gac	att	caa	cag	aca	gta	gga	ctt	agc	acc	caa	gaa	cca			2084
322	Asn	Ile	Val	Asp	Ile	Gln	Gln	Thr	Val	Gly	Leu	Ser	Thr	Gln	Glu	Pro			
323			165					170					175						
325	ggt	gat	gtc	ttc	acc	tat	gca	gaa	ttc	gat	ggc	atc	ctt	ggt	atg	gca			2132
326	Gly	Asp	Val	Phe	Thr	Tyr	Ala	Glu	Phe	Asp	Gly	Ile	Leu	Gly	Met	Ala			
327			180					185					190						
329	tac	cca	tcg	ctc	gcg	tca	gag	tac	tcg	ata	cct	gtg	ttt	gac	aac	atg			2180
330	Tyr	Pro	Ser	Leu	Ala	Ser	Glu	Tyr	Ser	Ile	Pro	Val	Phe	Asp	Asn	Met			
331		195				200					205								
333	atg	aac	cga	cac	cta	gta	gct	caa	gac	ttg	ttc	tcg	gtt	tac	atg	gac			2228
334	Met	Asn	Arg	His	Leu	Val	Ala	Gln	Asp	Leu	Phe	Ser	Val	Tyr	Met	Asp			
335	210				215						220				225				
337	agg	aat	ggc	cag	gag	agc	atg	ctc	acg	ctt</									

09/69 755

Missing mandatory <220> and  
<223> features to explain  
artificial sequence. See # 12  
on Error Summary Sheet.

<210> 4

<211> 390

<212> PRT

<213> Artificial Sequence

<400> 4

Met Asn Phe Leu Lys Ser Phe Pro Phe Tyr Ala Phe Leu Cys Phe Gly



## VERIFICATION SUMMARY

DATE: 08/30/2000

PATENT APPLICATION: US/09/643,755

TIME: 14:32:20

Input Set : A:\Sequence

Output Set: N:\CRF3\08302000\I643755.raw

L:14 M:270 C: Current Application Number differs, Replaced Application Number  
L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:429 M:258 W: Mandatory Feature missing, <220> FEATURE:  
L:429 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:

**NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES**

The nucleotide and/or amino acid sequence disclosure contained in this application does not comply with the requirements for such a disclosure as set forth in 37 CFR 1.821 - 1.825 for the following reason(s):

- ☐ 1. This application clearly fails to comply with the requirements of 37 CFR 1.821 - 1.825. Applicant's attention is directed to these regulations, published at 1114 OG 29, May 15, 1990 and at 55 FR 18230, May 1, 1990.
- ☐ 2. This application does not contain, as a separate part of the disclosure on paper copy, a "Sequence Listing" as required by 37 CFR 1.821(c).
- ☐ 3. A copy of the "Sequence Listing" in computer readable form has not been submitted as required by 37 CFR 1.821(e).
- ☒ 4. A copy of the "Sequence Listing" in computer readable form has been submitted. However, the content of the computer readable form does not comply with the requirements of 37 CFR 1.822 and/or 1.823, as indicated on the attached marked-up copy of the "Raw Sequence Listing."
- ☐ 5. The computer readable form that has been filed with this application has been found to be damaged and/or unreadable as indicated on the attached CRF Diskette Problem Report. A substitute computer readable form must be submitted as required by 37 CFR 1.825(d).
- ☐ 6. The paper copy of the "Sequence Listing" is not the same as the computer readable form of the "Sequence Listing" as required by 37 CFR 1.821(e).
- ☐ 7. Other: \_\_\_\_\_

**Applicant must provide:**

- ☒ An initial or substitute computer readable form (CRF) copy of the "Sequence Listing"
- ☐ An initial or substitute paper copy of the "Sequence Listing", as well as an amendment directing its entry into the specification
- ☒ A statement that the content of the paper and computer readable copies are the same and, where applicable, include no new matter, as required by 37 CFR 1.821(e) or 1.821(f) or 1.821(g) or 1.825(b) or 1.825(d)

For questions regarding compliance with these requirements, please contact:

For Rules Interpretation, call (703) 308-1123  
For CRF submission help, call (703) 308-4212  
For PatentIn software help, call (703) 308-6856

**Please return a copy of this notice with your response.**